

Research Commentary

Can Local Assessment Data be Successfully Used as Part of an Arizona A-F Accountability System?

by Sarah Estrada, Ph.D., Christine Burnham, Ph.D., Jason K. Feld, Ph.D.,
John Richard Bergan, Ph.D., and John Robert Bergan, Ph.D.
Assessment Technology Incorporated



Fundamental Questions: Assessment Technology, Incorporated (ATI) is in a unique position to assist the State Board in examining questions related to the feasibility of including a menu of local and/or statewide assessments in an Arizona A-F accountability system. ATI leadership staff members have extensive expertise and decades of experience in measurement, educational policy, and research approaches to education. As part of the research community, ATI staff members have authored a continuous stream of books, journal articles, white papers, and technical reports. As a company, ATI has conducted ongoing research for more than 25 years in our home state of Arizona and nationwide on topics related to educational measurement, assessment design, assessment implementation, and statistical analysis. Additional information about ATI's research is available on the Publications page of the ATI website (ati-online.com).

As of this writing, ATI currently serves 330 districts and charter schools throughout the state of Arizona. This group includes local educational agencies (LEA) from each Arizona County representing both rural and urban areas and serving students with a wide variety of demographic characteristics. These districts and charter schools use the Galileo® K-12 Online Instructional Improvement and Effectiveness System to design, administer, statistically analyze, and report on local assessments including using the resulting data to forecast statewide test performance at multiple points throughout the school year.

In response to the State Board's call for papers, this commentary summarizes results of ATI's ongoing research to help the State Board evaluate whether local assessment data can be successfully used as part of an A-F accountability system. Specifically, this commentary addresses two fundamental questions:

- To what extent do Arizona LEAs administer local assessments that can produce data for use in an accountability system?
- Are the data from existing local assessments sufficiently high-quality to be appropriate for use in an accountability system?

Extent of Arizona LEA Administration of Local Assessments: ATI conducts ongoing research regarding local assessment administration using Galileo. This research suggests that Arizona LEAs are already implementing a wide variety of local assessments. For example, in the 2014-15 school year, 294 Arizona districts and charter schools administered local district/charter-wide assessments using Galileo. As Table 1 illustrates, this group collectively administered 23,707 district/charter-wide assessments in math, English language arts, science, and writing in grades kindergarten through 12 resulting in 3,781,995 student scores. On average, each district or charter school administered 81 assessments across all grades and content areas.

TABLE 1

Local district/charter-wide assessment administration by Arizona LEAs using Galileo®

Local District/Charter-wide Assessment Administration by Arizona LEAs using Galileo		
Grade	Number of Assessments	Number of Student Scores
KG	855	94,252
1	1,437	235,622
2	1,981	335,100
3	2,200	375,464
4	2,494	414,965
5	2,311	367,320
6	2,437	381,535
7	2,420	384,306
8	2,523	400,460
9	845	150,662
10	3,109	518,300
11	620	77,093
12	475	46,916
Total	23,707	3,781,995

LEAs using Galileo administer local assessments to serve multiple purposes including guiding instruction, improving the effectiveness of instruction, and providing data for use in accountability initiatives. For example, many LEAs use data from local assessments to evaluate teachers and principals consistent with the Arizona Framework for Measuring Educator Effectiveness. LEAs also administer various types of assessments depending on their specific needs including ATI pre-built district/charter-wide assessments designed to mirror released statewide test blueprints and customized district/charter-wide assessments aligned to their own district/charter curriculum and pacing guide. Although Table 1 focuses on district/charter-wide assessments (e.g., pretests, posttests, benchmarks), it is worth noting that LEAs typically also administer a wide variety of school and classroom assessments (e.g., quizzes, unit tests, end-of-course exams). Assessments in the current investigation were administered both online and offline; however, LEAs are increasingly moving toward online assessment administration in part to facilitate the use of technology-enhanced items similar to those included on the new AzMERIT statewide assessments.

Overall, ATI's ongoing research suggests that the majority of Arizona LEAs already administer a wide variety of local assessments. This suggests that local assessments can potentially provide a wealth of data for use as part of the A-F accountability system.

Quality of Data Resulting from Local Assessments: In evaluating the feasibility of implementing a menu of assessments, the State Board will need to establish criteria for the inclusion of assessments and determine whether assessments meeting these criteria are available. The selected criteria should help ensure that assessments included in the menu will produce high-quality assessment data appropriate for use in an accountability system. Typically,

assessments implemented for purposes related to accountability are required to show evidence of alignment to a desired set of standards as well as adequate reliability and validity.

Local assessments administered by Arizona LEAs using Galileo® are composed of items from secure item banks created by ATI. These secure item banks contain multiple types of items including a wide variety of types of technology-enhanced items. All items are high-quality and aligned to Arizona standards via a multi-stage process including the development of item specifications, item construction, and item review with certification. This multi-stage process helps ensure that assessments created using these items are aligned to the desired standards.

LEAs using Galileo are also supported by ATI in the design and implementation of reliable local assessments that provide information about ability and standards mastery to guide instruction for students of a wide range of abilities. In contrast to statewide assessments which are administered over several hours, local district/charter-wide assessments are typically designed to be administered within a single class period. For this reason, local district/charter-wide assessments usually only include approximately 35 to 50 items. ATI research reveals that assessments composed of 35 to 50 items routinely display reliabilities at or above 0.80, with many assessments displaying reliabilities at or above 0.90. A report within the Galileo system provides LEAs with an estimate of the reliability of each local district/charter-wide assessment administered, supporting ongoing evaluation of the quality of the resulting data.

The predictive validity of local assessments is typically evaluated with respect to a relevant criterion measure such as the statewide assessment. For a number of years, Arizona LEAs have been required to administer statewide assessments in multiple grades and content areas, ensuring that LEAs have data from a criterion measure readily available to support this research. Once districts/charters have uploaded their statewide assessment data for individual students into Galileo, ATI conducts an investigation of the predictive validity of local district/charter-wide assessments. This ongoing predictive validity research has consistently indicated that student performance on local district/charter-wide assessments is adequately correlated with student performance on the statewide AIMS assessments. ATI research has also indicated that Galileo Risk Levels based on performance on multiple local district/charter-wide assessments effectively forecast student performance on the statewide AIMS assessments with approximately 85% accuracy.

With the recent release of data from the new statewide AzMERIT assessments, ATI has also initiated research to investigate the relationship between student performance on local district/charter-wide assessments and student performance on the AzMERIT assessments. In a preliminary investigation, correlation analyses were conducted for the first 100 districts and charter schools to provide ATI with their AzMERIT assessment data for individual students in grades three through eight in math and English language arts. Analyses for high school are currently underway. The sample included 226,625 paired sets of scores from local assessments administered within the 2014-15 school year and the statewide AzMERIT assessment in the same grade and content area administered in spring 2015. Student scores on local assessments reflected 1,674 district/charter-wide assessments. As Table 3 shows, the mean correlations range from 0.71 to 0.79 across grades and content areas with an overall mean across all grades and content areas of 0.75. The observed correlations suggest that student scores on the local assessments were adequately related to student scores on the statewide AzMERIT assessments.

TABLE 3
Correlation of scores from local assessments with scores from AzMERIT

Correlation of Scores from Local Assessments with Scores from AzMERIT			
Grade and Content Area		Number of Students	Mean Correlation
ELA	3	19,519	0.77
	4	19,572	0.77
	5	19,011	0.78
	6	19,186	0.79
	7	17,540	0.77
	8	17,928	0.77
	All Grades	112,756	0.78
Math	3	18,832	0.71
	4	19,581	0.75
	5	19,090	0.74
	6	19,147	0.75
	7	18,384	0.72
	8	18,835	0.71
	All Grades	113,869	0.73
All Grades and Content Areas		226,625	0.75

Overall, ATI's research indicates that local assessments administered by Arizona LEAs using Galileo® are aligned to Arizona standards and display adequate levels of reliability and validity to meet current LEA purposes. These findings suggest that data from many local assessments is likely to be of sufficient quality for use in an accountability system, depending on the specific criteria established by the State Board. In addition, the observed relationship between performance on local and statewide assessments suggest that the integration of data from multiple local and/or statewide assessments is a reasonable pursuit.

Conclusion: This commentary summarized the results of ATI ongoing research to help the State Board address two fundamental questions related to the feasibility of implementing a menu of local and/or statewide assessments in an A-F accountability system. First, to what extent do Arizona LEAs administer local assessments that can produce data for use in an accountability system? Second, are the data from existing local assessments sufficiently high-quality to be appropriate for use in an accountability system?

ATI research demonstrates that a large number of Arizona LEAs are already administering local assessments aligned to Arizona standards at multiple time points throughout the year, with many LEAs using the resulting data for other accountability purposes such as teacher and principal evaluation. Therefore, a large pool of local assessment data is potentially available for use as part of the A-F accountability system. To ensure the quality of data from local assessments, investigations of the reliability and validity of local assessments can easily be accomplished using existing statistical methods. In ATI research, local assessments administered by Arizona LEAs using Galileo are routinely found to be reliable and valid for

current LEA purposes, with performance on local assessments adequately correlated with and effectively forecasting performance on statewide assessments.

Overall, the current research suggests that, for the majority of Arizona LEAs, high-quality data from reliable and valid local and statewide assessments is readily available. This supports the feasibility of a menu of assessments, provided that assessments included in this menu meet criteria established by the State Board to ensure the quality of the resulting data. Although it is beyond the scope of the current commentary, a number of outstanding questions remain about the best way to incorporate data from multiple assessments within an Arizona A-F accountability system. ATI would welcome the opportunity to engage in additional discussions with the State Board and to conduct additional research relevant to this issue.